

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 27, 2003, 03:34:07 ; Search time 1693 seconds
(without alignments)
6189.297 Million cell updates/sec

Title: US-09-677-653A-50
perfect score: 3374
Sequence: 1 MGDAVGASORPHNRGTRNV.....GKTAARVRARRARRAARN 647
Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132
Minimum Match length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=x1h
-O=/CGN_1/YGAPO.spool/US0967753/runat_20022003_164315_1262/app.query.fasta_1.839
-DB=EST -QFM=fastap -SUFFIX=p2n.rst -MINNATCH=0.1 -LOOPCII=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGNN=15 -MODE=LOCAL
-OUTFILE=pto -NORM=ext -HEAVYSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0967753 @CGN_1_113_@unit_20022003_164315_1262 -NCPU=6 -ICPU=3
-NO_XLDPY -NO_MMAP -LARGEQUERY -NEG SCORE=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRESHOLD=0.05 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST: *
1: em_estba: *
2: em_lesthun: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: qb_est1: *
10: qb_est2: *
11: qb_htc: *
12: qb_est3: *
13: qb_est4: *
14: qb_est5: *
15: em_lestron: *
16: em_lestron: *
17: qb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_plnt: *
21: em_gss_vrt: *
22: em_gss_fut: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rdt: *

Result No.	Score	Query Match	Length	DB ID	Description
1	121.5	3.6	3775	11 BC028165	BC028165 Homo sapi
2	116	3.4	1535	13 BM552027	BM552027 AGENCOURT
3	112	3.3	900	12 BF167575	BF167575 61174283
4	111.5	3.3	1069	14 BM13440	BM13440 AGENCOURT
5	111.5	3.3	169	14 BQ712831	BQ712831 AGENCOURT
6	109.5	3.2	2128	13 BI082474	BI082474 60287785
7	109	3.2	3575	11 BC010575	BC010575 Homo Sapi
8	108.5	3.2	915	13 BI413164	BI413164 602990033
9	108.5	3.2	1776	14 BM264949	BM264949 AGENCOURT
10	108	3.2	861	12 BG168931	BG168931 60232081
11	107.5	3.2	869	9 AL578091	AL578091 AL578091
12	106.5	3.2	686	17 AG139016	AG139016 Pan trogl
13	106.5	3.2	820	13 BI650823	BI650823 603297718
14	106	3.1	959	17 CNS03M9N	AL250584 Tetraodon
15	106	3.1	983	14 BG049851	BQ649851 AGENCOURT
16	106	3.1	1060	13 BM49598	BM49598 AGENCOURT
17	106	3.1	1061	17 CNS03379	AL25774 Tetraodon
18	105.5	3.1	635	9 AI541596	AT541596 SD0277.5
19	105.5	3.1	783	13 BI652061	BI652061 603298942
20	105.5	3.1	1020	13 BM479488	BM479488 AGENCOURT
21	105.5	3.1	1294	13 BI409855	BI409855 602961973
22	105	3.1	655	13 BI244281	BJ244281 60244281
23	105	3.1	1067	12 BF573817	BF573817 602241914
24	105	3.1	1485	14 BM907789	BM907789 AGENCOURT
25	104.5	3.1	908	12 BE736099	BE736099 601305947
26	104.5	3.1	962	14 BG055611	BQ955611 AGENCOURT
27	104	3.1	891	14 BG670497	BO670497 AGENCOURT
28	104	3.1	973	12 BFF81727	BF981727 602306079
29	104	3.1	978	14 BQ714511	BQ714511 AGENCOURT
30	104	3.1	291	17 BH71003	BH71003 LIGM+972
31	103.5	3.1	679	13 BM406277	BM406277 ESF58004
32	103.5	3.1	830	14 BM85604	BM988604 2_A12_T3
33	103.5	3.1	963	14 BG899717	BQ899717 AGENCOURT
34	103.5	3.1	1067	12 BG168781	BG168781 602320073
35	103.5	3.1	1092	14 BD676763	BQ676763 AGENCOURT
36	103.5	3.1	1093	13 BM557460	BM557460 AGENCOURT
37	103	3.1	937	13 BM461960	BM461960 AGENCOURT
38	103	3.1	1031	17 CNS05IRX	AL339270 Tetradon
39	102.5	3.0	663	14 BO917203	BO917203 OHB20122
40	102.5	3.0	750	13 BI768562	BI768562 603057176
41	102.5	3.0	942	14 BQ930180	BQ930180 AGENCOURT
42	102.5	3.0	1060	13 BM424198	BM424198 AGENCOURT
43	102.5	3.0	187	13 BM546425	BM546425 AGENCOURT
44	102	3.0	1054	17 CNS03PVI	AL255159 Tetradon
45	102	3.0	1259	14 BG060839	BQ960839 AGENCOURT

ALIGNMENTS

RESULT 1
BC028165
LOCUS BC028165
DEFINITION Homo sapiens, Protocadherin 17, clone IMAGE:5245587, mRNA.
ACCESSION BC028165
VERSION BC028165.1 GI:20380902
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 3775)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (08-APR-2002) National Institutes of Health, Mammalian

Db	1427 CHACCCGCTCCGGTCAACAAAGGCTCTACCGTGCCTCAGGTGCACAGAACACATCCC	1486
Qy	421 rProserserArgArgLeuProMetProAlaSpPheGlyArgInThrValA	441
Db	1487 GCGAGAGTACCTGGCTGTG-CTCGCCAGATCCC--GACCTGGCCAG-----	1535
Qy	441 IaasnasnProLysIleGluglinserLeuLeuLysGluThrLeuGlyCystYLeuValH	461
Db	1536 -.-.AACGCCCGCATCCACTCTCATCCCTCCCTGCACACATGCCGACCTGTCATCT	1590
Qy	461 IserLysMetArg--AsnProValPheGlnIleuthrProLaserSerPheGlyAlaV	480
Db	1591 ACACCTATGIGCTGIGAATCCC-----AGAACGGGCCATCTAGGCCCTGC	1638
Qy	480 alserPheAsnAsnProGlyArgLysGluArgThrArgAspLeuProAspTyrrhlyleA	500
Db	1639 GCTCCTTAAC-----TTCGAGCAGACCAAAGCTTGTAGTCAAAGGTCTGCTA	1689
Qy	500 rGaspSer-----PheAspGlnAsnMetSerThrAlaValAlaHisPheA	515
Db	1690 AGGACTCGGGGGCGCCGCGCACCTGGAGAAGCACCGCACGGTGAGCTGCTAG	1749
Qy	515 rSerLeuSerHisSerCysSerSerIleValThrLysThrTyrglnGlyYrpGluGly---	533
Db	1750 ACCTGAATGACAAACGCCAGTGATCTGCTGCCACGGCTGAGAACGACACCGCGGAGC	1809
Qy	534 -----valThrAsnValAsnThrProHeG	542
Db	1810 TGCAGGTGGCGCAAGCTGGCTATCTGGTGAGCACTGCGCCCTAGACA	1869
Qy	542 IyglnPheAlaAlaAlaGlyLeuLeuLys-----AsnGluGluLe	556
Db	1870 GCGACTTGGCGAGAGCGGCCGCTCACCTACGAGATGTCGACGCGAACGAGACCAC	1929
Qy	556 eucyLeuAlaLaspAlaLeuAlaThrArgLeuThrGlyValTyroProAlaLaspAla-	575
Db	1930 TGTGAGATCCACGCCCTCCACGGGAGATCGCAGCTGACCCCTCTGGAGGACG	1989
Qy	576 -----PheAlaAlaAlaValSerAlaPheAlaAlaAlaAsnMetLeusers	590
Db	1990 TGCAGGCCGTTGGAGCTGGTGTGAGGTCACCGACCCACGCCCTACCTGTGCG	2049
Qy	590 erValLeuLysSerGluAlaLarSerIleLeuLysSerValGlyGluThrAlaValG	610
Db	2050 CTGGCCAAG-----CTCATCATCCGCTGGGTGAGG-----G	2082
Qy	610 lyAlaAlaGlnSerGlyLeuAlaLysLeuProGly	621
Db	2083 GATCCCTCCCCAGGGGTACCAACGGGTAAGGGC	2117
RESULT 2		
BM52027		
DEFINITION	BM52027 1535 bp mRNA linear EST 20-FEB-2002	
LOCUS	AGENCOURT_6543677 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5549404	
ACCESSION	5' mRNA sequence.	
VERSION	BM52027.1 GI:18789547	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	I (bases 1 to 1355)	
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/.	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Straubberg, Ph.D. Email: cgabsr@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agenourt Bioscience Corporation Clone distribution: MGC clone distribution information can be	
FEATURES		
SOURCE		
BASE COUNT	220 a	
ORIGIN	596 c	
Alignment Scores:	0.531	Length: 1535
Score:	116.00	Matches: 72
Percent Similarity:	30.28%	Conservative: 37
Best Local Similarity:	20.00%	Mismatches: 114
Query Match:	3.44%	Indels: 137
DB:	13	Gaps: 15
US-09-677-653A-50 (1-647) x BM552027 (1-1535)		
Qy	265 GlyGlyGlnTyraLeuThrProLeuProLysSerProLysAspValSerGluAlaTyR	284
Db	45 GGAGGTGAAGCTGGTGTGAGGCCAGTGGTGTCCGCCACCTCACCCAGGAGCTGGCTGG	104
Qy	285 AlaLeuAlaThrLeuThrPheAlaLaaGProSerSerAlaAlaAlaLeuAlaPheValTrp	304
Db	105 GGC-----ASGACCGGAGGGAGCACGCTGCGCTCCGCGCCCTGCG	149
Qy	305 AlaGlyLeuProGlnGlyGlyThrAlaProAlaGly--ThrProAla--TrpGluIn	322
Db	150 -----CTTCCGAGCTAGGGTGGACGCCAGCTCTGCCATGTTGAGGAA	203
Qy	323 AlaSerSerGlyGly-TyrLeuThrTrpArgHisAsnGlyThrThrPheProAlaGly	342
Db	204 GGGAAAGGAGGGGCTC-----	222
Qy	342 rValSerSerGlyGly-TyrLeuThrTrpArgHisAsnGlyThrThrPheProAlaGly	362
Db	223 -----AlaSerSerGlyGly-TyrLeuThrTrpArgHisAsnGlyThrThrPheProAlaGly	245
Qy	362 rTrpThrAspPheAlaSerAlaGlyAspThrValThrPhe-----	375
Db	246 CTTGGTGGAGAGTAGTCAGTTAGGGGCCCTCCCTTTTTCGGGGGGGACCC	305
Qy	376 -ArgGlnAlaAlaValAspGluValValValThrAsnAsnProAlaGlyGlyGlySerAl	395
Db	306 CCCCAACCCCCAACCGCACCTCTCCCCGAAATCCCGCCGGGGGGGGGGCG	365
Qy	395 aProThrPheThrValArgValPro--ProSer-----AsnAlaTyThr-----	409
Db	366 TCCCTCCCTCCCTTCATCCGCAATTGTGACTCCGGCCCTCACATACA	425
Qy	409 -----AsnThrValPheArgAsnThrLeuLeuGluThrArgProSerSerArgArgLe	409
Db	426 CGCCATGCCCCCTTGCTGCAACCTCACCGCAATTCTCTCCTCCCCCTCAC	485
Qy	410 -----AsnThrValPheArgAsnThrLeuLeuGluThrArgProSerSerArgArgLe	427
Db	486 TTTCCTGGCTGCCCCCTGCTCACACTGCCTCTGTCCTGGCCGCC	545
Qy	427 uGluLeuPro-----	430
Db	546 CTCGGCCCTTACGGCTCTTGTGCTCTTGTGCTCTTGTGCTCTGCGCCCC	605

found through the I.M.A.G.E. Consortium/LNLL at:
<http://image.llnl.gov>
 Plate: LLNL260 row: a column: 17
 High quality sequence start: 16
 High quality sequence stop: 269.

Location/Qualifiers
 1..1535
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5549464"
 /clone.lib="NIH MGC 71"
 /tissue_type="leiomyosarcoma"
 /note="Ovarian: uterus; Vector: pCMV-SPOR6; Site_1: NotI; Site_2: SAI1; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 2.1 kb.
 5 others

QY	285	-----	AlateUhishrLeu-ThrPheAlaArgP	294
Db	472	CACAGGAAAGGGGGACACAACGGGGGGCAATCCACCTCTTCATCCCCCGGCC	531	
QY	294	oSerSerAlaAlaAlaLeuAlaPheValTrpAlaCylLeuProGlnGlyLysThrAlaP	314	
Db	532	AGCTACTACTGCCCTCATC-----ATTCCACGGCTGGTCTAAGCC	573	
QY	314	oAlaGlyThrProAlaTrpLysGluAlaSerSerGlyGlyTyriLeuThrPheAlaArgP	333	
Db	574	CCCATCCTCCCTGCCACCATATCACGPGCAGCAATCCATTGGAATCCCTCGGGCACACT	693	
QY	334	--AsnGlyThrThrPheProAlaGlySerValSerValSerValSerValLeuProGlyLysThrAlaP	352	
Db	634	ACCCACGCGATGTCACCCACGPGCAGCAATCCATTGGAATCCCTCGGGCACACT	693	
QY	352	aLeuGluArgTyrAspProLysAspGlyLysSerTrpThrAspHeAlaSerAlaGlyAspTh	372	
Db	694	CTTAAACCACTT-----	720	
QY	372	rValThrPheArgGlnVal-----AlaValAspGluValValValThrAs	387	
Db	721	ATGGCGGACGCCGCTGGTACACACTAGAACGACTTATACCGGCATATCGGATA	780	
QY	387	nAsnProAlaGlyGly-----SerAlaProThrPheThrValArgValProPro	405	
Db	781	TTACCTACTTAGCGGTATCCCTATCCCGATTAATGTCACAATTCTACCTCCACCCCG	840	
QY	405	r-----AsnAlaTyrrThrAsnThrValPheAlaGlyAsnThrLeuG	419	
Db	841	CCCTAACATAGTCACCTGGGGCACTCACCTCCACATACTGTCCTACAGCAACGA	900	
QY	419	uthrargPro-----SerSerArgAlaGlyLeuProMetProProAlaAspP	436	
Db	901	GGCAAACGCCAACCTTACACCAATCCATACCCCTTGATCCCGCCCT-----	955	
QY	436	eGlyGlnThrValAlaAsnAsnProLysLeuGlnSer-----LeuLeuLysGlu	454	
Db	956	---CACACCAACCAACCTAACCTAACACTAACACCATTGCGACCACTACTCAATAA	1011	
QY	454	rLeuGlyCystyriLeuValHisSerLysMetArgAsnProValPheGlnLeuThrProA	474	
Db	1012	CCTAACCATACAC-----CACACCTCTAACCCACAACCTCTCACCCCTACTCCACC	1065	
QY	474	aserSerPheGlyAlaValserPheAsn-----	485	
Db	1066	ACATCCTTAAACGCTTGCCACATTAATCTACCCAAAGCTCTTATCCCGCGTGACACGGTAC	1125	
QY	485	OGlyTyrglurArgThrArgAspLeuProAspPheThrGlyIleArgAspSerPheAspG	505	
Db	1126	AGCCCTACCCAGACACCGATATCCCCCGGCCACGGTGTACAACGGACAAACG	1185	
QY	505	nAsnMetSerThrAlaValAlaLysPheArgSerLeuSerHisSerCysSerIleValTh	525	
Db	1186	ATGACTAACACCA-----CCCTCTCATCTAACCCATCC-----	1222	
QY	525	rLysThrTygIglngly	530	
Db	1223	-----ACCTACTACGGC	1234	
RESULT	7			
BC010575	BC010575	Homo sapiens, Similar to E74-like factor 1, clone IMAGE:4338775, mRNA	3572	bp mRNA linear HTC 12-JUL-2001
LOCUS				
DEFINITION				
ACCESSION	BC010575			
VERSION	BC010575.1	GI:14714847		
KEYWORDS	HTC.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;				
REFERENCE	1	(bases 1 to 3572)		
AUTHORS	Strausberg, R.			
TITLE	Direct Submission			
JOURNAL				
REMARK				
COMMENT				
CDNA Library	Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
DNA Sequencing	by: Baylor College of Medicine Human Genome			
Sequencing Center				
Center code:	BCM-HSC			
Web site:	http://www.hgsc.bcm.tmc.edu/cdna/			
Contact:	villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.			
Clone distribution:	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: http://image.llnl.gov			
Series:	IRAK plate: 12 Row: 0 Column: 13			
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenomeScan gene prediction, Similarity but not identity to Protein				
This clone has the following Problem: incomplete processing				
FEATURES	Location/Qualifiers			
Source	1. .3572			
/organism="Homo sapiens"				
/db_xref="taxon:9606"				
/clone="IMAGE:338775"				
/tissue_type="Ymph, lymphoma"				
/clone_id="NIM_MGC_85"				
/note="Vector: PCMV-SPORR6"				
BASE COUNT	1135 a 678 c 703 g 1056 t			
ORIGIN				
Alignment Scores:				
Pred. No.:	12.6	length: 3572		
Score:	10.9.00	Matches: 117		
Percent Similarity:	35.22%	Conservative: 82		
Best Local Similarity:	20.71%	Mismatches: 214		
Query Match:	3.23%	Indels: 153		
DB:	11	Gaps: 22		
US-09-677-653A-50 (1-647) x BC010575 (1-3572)				
QY	97	LysTyriLeuAspProAlaGlyAlaThr-----GluSerAlaArgAlaValGlyGlu	113	
Db	781	AAATACCGACTCACGGGAGCCATACCAAGAGCAAGCTTAAGNGAAAAGGAAGA	840	
QY	114	TyrSerLysIleProAspGlyIleValValAspSerValAspAlaGluIleArgGlu	133	
Db	841	AAAATCAACGACCA-----	855	
QY	134	TyrAspGluGlyCysProValAlaThr-----AspValSerValPro-----LeuAsp	149	
Db	856	--CGAACGGAACACAAATTATCTTGGGAGTT-----TTA	948	
QY	150	GlyArgGlnIrpSerIleUserIlePhePheProMetPheArgThrAlaValAla	169	
Db	913	GGAAGGGAAACACAAATTATCTTGGGAGTT-----TTA	942	
QY	170	ValAlaAsnValGluAsnLysGluMetSerLeuAspValAlaAspLeuIleGluTrp	189	
Db	949	CGGGACTGCTCCAGGACAAGCT-----ACTTGCTTAATACTACATCAGTGG	996	

/tissue_type="hypernephroma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site:1: NotI;
 Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

	BASE COUNT	ALIGNMENT SCORES:	LENGTH:
Db	345	108.00	861
Db	373	36.40%	71
Db	393	35.40%	28
Db	329	36.10%	93
Qy	413	108.00	861
Db	365	36.10%	80
Qy	432	36.10%	16
Db	425	36.10%	12
Qy	451	108.00	861
Db	485	36.10%	277
Qy	471	36.10%	167
Db	533	36.10%	296
Qy	491	36.10%	209
Db	592	36.10%	314
Qy	510	36.10%	248
Db	647	36.10%	334
Qy	525	36.10%	365
Db	707	36.10%	305
RESULT 10	BG168931	108.00	861
DEFINITION	861 bp mRNA linear EST 06-FEB-2001	108.00	861
ACCESSION	BG168931	108.00	861
VERSION	BG168931.1 GI:12675634	108.00	861
SOURCE	human.	108.00	861
ORGANISM	Homo sapiens	108.00	861
Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	108.00	861	
REFERENCE	I (bases 1 to 861)	108.00	861
AUTHORS	NIH-MGC http://mgc.ncbi.nih.gov/	108.00	861
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	108.00	861
JOURNAL	Unpublished (1999)	108.00	861
COMMENT	Contact: Robert Strausberg, Ph.D.	108.00	861
TISSUE	Procurement: ATCC	108.00	861
CODNA	Library Preparation: Life Technologies, Inc.	108.00	861
LIBRARY	Arrayed by: The I.M.A.G.E. Consortium (LLNL)	108.00	861
DNA	Sequencing by: Incyte Genomics, Inc.	108.00	861
Clones	DNA distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://Image.llnl.gov	108.00	861
Plate	LIAM0144 row: b column: 21	108.00	861
High quality sequence start: 2.	High quality sequence stop: 772.	108.00	861
FEATURES	Location/Qualifiers	108.00	861
source	1. organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4415156" /clone_id="NIH_MGC_89"	108.00	861

/tissue_type="hypernephroma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site:1: NotI;
 Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

	BASE COUNT	ALIGNMENT SCORES:	LENGTH:
ORIGIN	284	108.00	861
QY	263	108.00	861
Db	108	108.00	861
Qy	278	108.00	861
Db	168	108.00	861
Qy	297	108.00	861
Db	210	108.00	861
Qy	315	108.00	861
Db	249	108.00	861
Qy	335	108.00	861
Db	306	108.00	861
Qy	347	108.00	861
Db	366	108.00	861
Qy	361	108.00	861
Db	426	108.00	861
Qy	362	108.00	861
Db	486	108.00	861
Qy	376	108.00	861
Db	546	108.00	861
Qy	396	108.00	861
Db	606	108.00	861
Qy	415	108.00	861
Db	666	108.00	861
Qy	435	108.00	861
Db	723	108.00	861
Qy	453	108.00	861
Db	783	108.00	861
Qy	466	108.00	861

/tissue_type="hypernephroma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site:1: NotI;
 Site:2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

	BASE COUNT	ALIGNMENT SCORES:	LENGTH:
ORIGIN	284	108.00	861
QY	263	108.00	861
Db	108	108.00	861
Qy	278	108.00	861
Db	168	108.00	861
Qy	297	108.00	861
Db	210	108.00	861
Qy	315	108.00	861
Db	249	108.00	861
Qy	335	108.00	861
Db	306	108.00	861
Qy	347	108.00	861
Db	366	108.00	861
Qy	361	108.00	861
Db	426	108.00	861
Qy	362	108.00	861
Db	486	108.00	861
Qy	376	108.00	861
Db	546	108.00	861
Qy	396	108.00	861
Db	606	108.00	861
Qy	415	108.00	861
Db	666	108.00	861
Qy	435	108.00	861
Db	723	108.00	861
Qy	453	108.00	861
Db	783	108.00	861
Qy	466	108.00	861

	VERSION	B1650823.1 GI:15565059
	KEYWORDS	EST, house mouse
PRED. NO.:	SOURCE	mus musculus
SCORE:	ORGANISM	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
PERCENT SIMILARITY:	REFERENCE	1 (bases 1 to 820) NIH-MGC http://mgc.nci.nih.gov/
BEST LOCAL SIMILARITY:	AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
QUERY MATCH:	JOURNAL	NIH-MGC http://mgc.nci.nih.gov/
US-09-677-653A-50 (1-647) x AG3139016 (1-686)	COMMENT	Unpublished (1999)
QY 242 ArgLeuThrTyrLys----AlaIleThrCysGluAlaAsnMetProThrLeuValAsp 259	CONTACT	Email: cgabbs-r@mail.nih.gov
Db 12 CGACHTCACTACAGGCTAGCCGCTAGCATGACTCGAGCTTGACTTCCTTGACCA 71	TISSUE PROCUREMENT	Robert Strausberg, Ph.D.
QY 260 GlnGlyPheTrp-----IleGlyGlyGlnTyAlaLeuThr 271	PREPARATION	Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
Db 72 GGAAGGTCTGACCAGGGCTTGTCTCCAGGCAGTCCTCACCAAGCCTGGACAA 131	CDNA LIBRARY	Life Technologies, Inc.
Y 272 ProThrSerLeuProGlnTrpAspValSerGluAlaTyrosineAlaLeuHisThrLeuThrPhe 291	CDNA LIBRARY	Arrayed by: The I.M.A.G.E. Consortium (LNNL)
Db 132 GCCCAAGGCCCTGCCCTCTTGTCTTCACAGCATAT----- 170	DNA SEQUENCING	DNA Sequencing by: Incyte Genomics, Inc.
QY 292 AlaArgProSerSerAlaAlaAlaLeuAlaPheValAlaGlyLeuProGlnGlyGly 311	Clone Distribution	MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNNL at: http://image.llnl.gov
Db 171 ----- 171	Plate:	LLAM11859 row: c column: 06
QY 312 Thr-AlaProAlaGlyThrPheProAlaTrpGluGlnAlaSerSerGlyGlyTyLeuThrT 331	High Quality Sequence	High quality sequence stop: 729.
Db 201 ACTTGCTCTCTGCT-----CGGATTGG-----TCTT 227	Location/Qualifiers	Location/Qualifiers
QY 331 RPAQHISAnGlySerValSerIleTyValLeuProGluGlyP 351	1. . 820	1. . 820
Db 228 GGAAAGAGAAAT-----CCCTGGCTTCCCTGGCTTCCT-----GGCC 266	/organism="Mus musculus"	/organism="Mus musculus"
QY 351 heAlaLeuGluArgTyrasProAsnAspGlySerItpThrAspPheAlaSerAlaGlyA 371	/db_xref="taxon:10090"	/db_xref="taxon:10090"
Db 267 TCCACCCCCCAAAATCAAACCCGGGGAACGCTTGTGAAACCCGGAAATCTGGC--- 322	/clone="IMAGE:5338301"	/clone="IMAGE:5338301"
QY 371 sphrValThrPheArgGlnValAlaValAspGluValValThrAsnAlaProAla 391	/tissue="Site-1: Sali"	/tissue="Site-1: Sali"
Db 323 ----- 338	/tissue="Site-2: Cloned unidirectionally. Primer: Oligo dT."	/tissue="Site-2: Cloned unidirectionally. Primer: Oligo dT"
QY 391 LyGlyLySarAlaProThrPheThrValArgValProProSerAsnAlaTyrosineAsnT 411	/dev_stage="10 months"	/dev_stage="10 months"
Db 339 CGGGGGCTCCCCAAAACACACCGGAGGCCACCGATCTAACGCCCTCGAAC 398	/lab_host="DH10B"	/lab_host="DH10B"
QY 411 hrval----- 417	/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Sali; Site_2: Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."	/notes="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Sali; Site_2: Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."
BASE COUNT	122 a 289 c 249 g 160 t	1. . 820
ORIGIN		
Alignment Scores:		
PRED. NO.:	Length:	820
SCORE:	Matches:	47
PERCENT SIMILARITY:	Conservative:	19
BEST LOCAL SIMILARITY:	Mismatches:	61
QUERY MATCH:	Indels:	25
DB:	Gaps:	8
US-09-677-653A-50 (1-647) x B1650823 (1-820)		
QY 417 ArgLeuGlyValLeuArgProThrTyrasPheValProAspProThrLeuGly----- 233		
Db 29 CGCTGCTACCTGAGCACGCGTCTGAGCTCCCTCTCACTCTGCGGGTCTCTCTC 88		
QY 459 TTCCAAATTACACGCGCCGAAATTCACAAGATTCCTCCGCT----- 502		
QY 436 heGlyGlnThrValAlaAsnAsnProLysThrLeuLeuGluGlnSerLeuLeuGluThrLeuG 456		
Db 503 ----- 527		
QY 456 lyCysTyroLeuValHisSerLysMetArgAsnProValPheClnLeuThrProAlaSerA 476		
Db 528 GATGCTACAC----- 557		
RESULT 13		
QY 476 erpheGlyAlaValSerPheAsnAsnPro 485		
Db 558 CCCACCTATCAAATTCTCTCAGCACCG 586		
B1650823 820 bp mRNA linear EST 12-SEP-2001		
DEFINITION 603297718F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5338301 5'		
ACCESSION B1650823		

QY	319	AlatrpGluGluAlaLaserSerSlyGlyTyreLeuThrPargHisAsnGlyThrThrPhe	338	QY	199	ValValAspSerGluGlnTripleAsnProThrAspThrThrTyryValArgIle	218
Db	371	GCTGGGGGACCCGGGGCACCGGGCAGCTGTCTGGTCCAGCAGTGGGCAGTATG	430	Db	835	ACGGTCACAGGGAG-----	::::TG
QY	339	ProAlaGlySerValSerTyryValLeuProGlyU	350	QY	219	ArgValLeu-AlgProThrTyryAspValPro-----AspProThrGluGly	233
Db	431	CGAGCACGGAGC---TCATAACCGCTGCCGAGGGG	463	Db	817	AAAGTAGCGAGCCCCGAGGGAGACGCCAACAGCGGCTCTGGCAGGGAGCGCT	758
RESULT	14			QY	234	-----LeuValArgThrValSe	239
CNS03M9N/c	LOCUS	CNS03M9N	959 bp DNA linear	QY	757	GGAGGGCTCTCTCTTGTCGACAGGGCTGAGCGGGAGCTGAAGACCTACAG	698
DEFINITION	Tetraodon nigroviridis genome survey sequence T7 end of clone	037K22 of library G from Tetraodon nigroviridis, genomic survey	sequence.	QY	239	rasp-----TyrragLethrThryLysAlaLethrThrySglalaAsnMetBroth	256
ACCESSION	AL250484	AL250484.1	GI:7971496	Db	697	AGACATCATCGCTCTCAACCCCTCTACAAAGTCAGTCAGTCCTGGCCCTCGTCACACCCAC	638
VERSION				QY	256	rueuValAspGlnGlyPheTrpIleGlyGlyLysValLeuThrProinSerIeupr	276
KEYWORDS	GSS; genome survey sequence.	Tetraodon nigroviridis.		Db	612	-----CAGCGCCTGGAGGGAGCACCCCGCT	581
ORGANISM	Tetraodon nigroviridis			Db	637	GGCG-----GCGTCACACAGACGCGCT--	613
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.			QY	296	rAlaAlaAlaLeuAlaPheValTrpAlaGlyIeuProGlnGlyGlyThrAlaProLgI	316
AUTHORS	Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.			Db	580	CTGGTCGCGCTGACGGCTCTGATCAGACTCAGCCGCTCGNGTACCTCG	521
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA Sequence			QY	316	yThPro----AlatrpGluGlnAlaSerSlyGlyTyreLeuThrPargHisa	334
JOURNAL	Unpublished			Db	520	GRACACGAGTGGGCGAGGCCACGTCGCGCTGAGTCGAGCAGAC	461
REFERENCE	2 (bases 1 to 959)			QY	334	nGlyThrThrPheProAlaGlySerValSerTyryValLeuProGluGlyPheAlaLeugI	354
AUTHORS	Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W. and Weissenbach,J.			Db	460	CGAGCGCTGTCACCTCTGGCCAGAGGAGCTGGC-----TCACACTGCT	410
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis			QY	354	uargTyrsProAspAsp-----GlyserPrp	364
JOURNAL	Unpublished			Db	409	CTGTGGCTCCAGGGAGTCGGCTCCAGATGATGGCAGGGCTGGTAGAGGGTGGGAC	350
REFERENCE	3 (bases 1 to 959)			QY	364	raspPheAlaSerAlaLysAspThrValThrPheArgLysValAlaLysValAspLysVal	384
AUTHORS	Genoscope			Db	349	A-----ACCCCATCTACCTCAGACATGGGAGCAGCTGACGGAGCTCC	299
TITLE	Direct Submission			QY	384	IvalThRAsnAsnProAla-GlyGlyLysSerAlaProThrPheThrValArgValPro	404
JOURNAL	Submitted (12-APR-2000)			Db	298	CACGAGCTCCAGACATCTGAGGGGGAGCTGAGTCAGTGAGCTGCCCCGGRC	239
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/~tetraodon .			QY	404	roSerAsnAlaTyryThrAsnThrValPheArgAsnThrLeuIgLuThrArgProSer	424
FEATURES	source			Db	238	CCCCCAGG-----CTGCCTCGCCCTGTATG	212
/organism="Tetraodon nigroviridis"				QY	424	erArgArgLeuGluleuProMetProProAlaAspPheGlyLysValAlaAsn	444
/Db_xref="taxon:9983"				Db	211	GGACCCGGTCCCTCCAGATCCCAGCCTCTAACAGGCCCTGTCCTGCTGAAGA	152
/clone_id="G"				QY	444	rolysIleGluLnsSerIleLeuLysGluThrLeuGlyCysTyryValHisSerLysM	464
/note="Genoscope sequence ID : C0BG037BF11LP1-end : T7"				Db	151	AGGAGTCGACGTCGACGAGCTGCAAGCCCTGGCAGAGGCAACATCCGGC	92
BASE COUNT	167 a	332 g	175 t	QY	464	erArgAsnPro	467
ORIGIN	284 c	1 others		Db	91	CCTCCAGGCC	81
Alignment Scores:				RESULT	15		
Pred. No.:	2.75	Length:	959	B0649851			
Score:	106.00	Matches:	71	LOCUS	B0649851	983 bp mRNA linear	EST 15-JUL-2002
Percent Similarity:	34.49%	Conservative:	48	DEFINITION	AGENCR-8341749 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6267739		
Best local Similarity:	20.58%	Mismatches:	138	5' mRNA sequence.			
Query Match:	3.14%	Indels:	88	VERSION	B0649851		
Db:	17	Gaps:	11	KEYWORDS	B0649851.1	GI:21774023	EST.
US-09-677-653A-50 (1-647) x CNS03M9N (1-959)							
QY	159	SerPheProMetPheArgGlyAlaAlaTyryValAlaLysValGluAsnLysGluMet	178				
Db	955	ACCTTGACCTCGCTTCCTCCAGGTTAGAAGCGGACTTGAGCTCCAGCTCG	896				
QY	179	SerLeuAspValAlaValAsnAspLeuIleGluTrpLeuAsnAsnLeuAlaAspThrPhe	198				
Db	895	TCTTGCCATGCTCCACGTCCTGATGGTGGAGAGAACATCCACCAGGAGTC	836				

SOURCE	human.	Qy	144 ValSerValProLeuAspGlyArgLntPSerLeuSerIlePheSerThrProMetPhe
ORGANISM	<i>Homo sapiens</i>	Db	111 GTGCAGAAGAGAGCGATCCGACCCCTCTGACAGCNG----- 314
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		Db	279 GTGCAGAAGAGAGCGATCCGACCCCTCTGACAGCNG----- 314
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		Qy	164 ArgThrAlaTyrValAlaValAlaAsnValGluasnLysGluMetSerLeuAspValVal 183
REFERENCE	National Institutes of Health, Mammalian Gene Collection (MGC)	Db	315 ----- 326
AUTHORS	Unpublished (1999)	Qy	184 AsnAspLeuIleGluTrpLeuAsnAsnLeuAlaAsnLeuAspTrpArgTrpValValAspGlu 203
JOURNAL	Contact: Robert Strausberg, Ph.D.	Qy	327 GATCAGATG-----TGGCTGCCAGTGGGAGGACTTGCGC----- 362
COMMENT	Email: cgaps-r@mail.nih.gov	Qy	204 GluTrpIleLeuPhThrIleAspPheSerAspThrTyr----- 214
TISSUE Procurement:	CGAP (Stanford)	Db	363 -----CTCAATGAGCGGCCACTATGGGGCTAACGGGTCTCAATAAGCA 407
cDNA Library Preparation:	Rubin Laboratory	Db	368 -----TGGCTGCCAGTGGGAGGACTTGCGC----- 476
DNA Sequencing by:	The I.M.A.G.E. Consortium (LLNL)	Qy	215 -----TyrValAlaGlearValLeuArgProThrTyrAspVal 227
Clone distribution:	NIH-MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:	Qy	408 GAAACTTGCTCAAAGCATGCTGTAGGCCAGGTGAGATCTGGAGGCCCTCTANGATGTC 467
Plate:	LICM2438 row: 1 column: 20	Qy	228 ProAspProThrGluGlyLeuValArgThrValSerAspThrArgLeuThrTyrLysAla 247
High quality sequence stop:	622.	Db	468 CCACCACCT----- 476
FEATURES	Location/Qualifiers	Qy	248 IleThrCysGluIlaAsnMetProThrLeuValAspGlyInlyPhePhePheIleGly----- 265
source	I. .983	Db	477 -----CCGATGGGCCGACCATCTCTACGCAACATCATG 515
/organism="Homo sapiens"		Qy	26 -----GlyGlnTyrAla---LeuThrProThrSerIleProGlnTyrAspValSerIle 282
/db_xref="taxon:9606"		Db	516 AAGGGATCGAGGTATGAGACCTCACAGAGATCAGCTACCCCTCTGAG----- 566
/clone="IMAGE:6267739"		Qy	283 AlatYralaleuHistrIleLeuThrPheAlaArgProSerSerAlaAlaAlaLeuAlaPhe 302
/tissue_type="hepatocellular carcinoma, cell line"		Db	567 -----AGTCGAGAGGATCATTTGCCAGACTCGCCCTTC 602
/lab_host="DH10B (phage-resistant)"		Qy	303 -----ValTrpAlaGlyLeuProGlnGlyGlyThralProAlaGlyThrPro 318
/note="Organ: Liver; Vector: pORB; Site_1: XbaI; Site_2:		Db	603 TGGATGAGAAATAGTTCCCAGATCAGGAAGGGAAACGCTACTGATTCAGCCCA 662
Ecoli; cDNA made by oligo-dT priming. Directionally cloned		Qy	319 AlaTrpGluIle-----AlaSerSerGlyGlyIleLeuThrPheGly----- 333
into EcoRI/XbaI sites using the following 5' adaptor:		Db	663 ---TGCACATAGCTCGGGGACATTGTCAGCATCTGGGGCTCTGAGCT 719
GCAACGAG(G). Size-selected >500bp for average insert size		Qy	334 AsnGlyIleThrPheProAlaGlySerValSerIleLeuProGlyLeuAlaPhe 353
1.8kb. Library constructed by Ling Hong in the laboratory		Db	720 CATGGGACTGAC---CCCTCGGACTCTGGGAATCCCATTTGCTATGGATGCGACCAACAG 776
of Gerald M. Rubin (University of California, Berkeley)		Qy	354 GluIgTyrAspProAsnAspGlySerIlePheSerPheAsp-----PheAlaSerAlaGly 370
using ZAP-cDNA synthesis kit (Stratagene) and Superscript		Db	777 GAAGCTTAAGGCCATCAGGTTCTGGGGCATAAAAGACGTTGCCAAAGCCCTG 836
TII RT (Life Technologies). Note: this is a NIH-MGC		Qy	371 AspThrValThrPheArgGlnValAlaValAspIleuValValThrAsnProAla 390
Library."		Db	837 GAACTTGCGCCGCCCCGGCA-----AGGCCAA 869
BASE COUNT	219 a 291 c 304 g 169 t	Qy	391 GlyGly-GlySerAlaProThrPheThrVal--ArgValProProSerAsnAla 407
ORIGIN		Db	870 AATGGAAGGCCGCGGGAGGATACTCTCCCGGAAGCCCGCTCGGG 924
Alignment Scores:			Search completed: February 27, 2003, 06:39:07
Pred. No.:	2.86	Length:	Job time : 1718 secs
Score:	106.00	Matches:	
Percent Similarity:	33.41%	Conservative:	
Best Local Similarity:	22.20%	Mismatches:	
Query Match:	3.14%	Indels:	
DB:	14	Gaps:	
US-09-677-653A-50 (1-647) x DQ649851 (1-983)			
Qy	29 ValAsnGlyArgArgGlnArgArgArgGlyArgGlnValSerProProAspAsn	Db	603 TGGATGAGAAATAGTTCCCAGATCAGGAAGGGAAACGCTACTGATTCAGCCCA 662
Qy	::::::: :::::	Qy	319 AlaTrpGluIle-----AlaSerSerGlyGlyIleLeuThrPheGly----- 333
Db	10 CTGAGCGGGTCCGAGCCGCCGAGGCCGACGGGGCGGGCTACTCCG----- 60	Db	663 ---TGCACATAGCTCGGGGACATTGTCAGCATCTGGGGCTCTGAGCT 719
Qy	49 PheThrAlaAlaAlaGlnAspLeuAlaGlnInserIleAspAlaAsnThrValThrPhePro	Qy	334 AsnGlyIleThrPheProAlaGlySerValSerIleLeuProGlyLeuAlaPhe 353
Db	61 -----GAACTTGCGCCGCCCCGGCA-----AGGCCAA 869	Db	720 CATGGGACTGAC---CCCTCGGACTCTGGGAATCCCATTTGCTATGGATGCGACCAACAG 776
Qy	69 AlaAsnIleSerSerMetProGluPheArgGlnTrp-----AlaLyGly	Qy	354 GluIgTyrAspProAsnAspGlySerIlePheSerPheAsp-----PheAlaSerAlaGly 370
Db	94 AGCCGGCCGCCATGGCCGCT-----ACAAACGCGTGTATCGGCACGGGGAGCG 147	Db	777 GAAGCTTAAGGCCATCAGGTTCTGGGGCATAAAAGACGTTGCCAAAGCCCTG 836
Qy	84 LysIleAspLeuAspSerAspSerIleGlyTrpPheLysTyrLeuAspProAlaGly	Qy	371 AspThrValThrPheArgGlnValAlaValAspIleuValValThrAsnProAla 390
Db	148 CATGG-AACCTGGAGAACCGCTCAGGGCTGAGGCCGACCTGAGCCGGCGGG 206	Db	837 GAACTTGCGCCGCCCCGGCA-----AGGCCAA 869
Qy	104 AlaThrGluSerAlaArgAlaIvalAlaGlyGluTyrsSerLysIleProAspAlaLeuVal 123	Qy	391 GlyGly-GlySerAlaProThrPheThrVal--ArgValProProSerAsnAla 407
Db	207 CACGAGGAGGCCAGAGCGCC--GGCCAG-----GCCTACGAGATCTGGCTATGAG 257	Db	870 AATGGAAGGCCGCGGGAGGATACTCTCCCGGAAGCCCGCTCGGG 924
Qy	124 PhesValAlaGluIleArgGluIleTyAsnGluGluCysProValValThrAsp 143		
Db	258 TTGACATC-----TGCTTCACCTCA 278		